

REVISED UG SYLLABUS UNDER CBCS
(Implemented from Academic Year 2020-21)
PROGRAMME: FOUR YEAR B.Sc.,

Domain Subject: STATISTICS

Skill Enhancement Courses (SECs) for Semester V, from 2022-23
(Syllabus-Curriculum)

Structure of SECs for Semester – V

(To choose one pair from the three alternate pairs of SECs)

Univ Code	Cours es 6 & 7	Name of Course	Th. Hrs. / Week	IE Mar - ks	EE Mar - ks	Credit s	Prac .Hrs ./ Wk	Mar - ks	Credit s
	6A	Operations Research - I	3	25	75	3	3	50	2
	7A	Operations Research – II	3	25	75	3	3	50	2

OR

	6B	Statistical Process and Quality Control	3	25	75	3	3	50	2
	7B	Computational Techniques and R Programming	3	25	75	3	3	50	2

OR

	6C	Econometrics	3	25	75	3	3	50	2
	7C	Regression Analysis	3	25	75	3	3	50	2

Note-1: For Semester–V, for the domain subject Statistics, any one of the three pairs of SECs shall be chosen courses 6 and 7, i.e., 6A & 7A or 6B & 7B or 6C & 7C. The pair shall not be broken (ABC allotment is random, not on any priority basis).

Note-2: One of the main objectives of Skill Enhancement Courses (SEC) is to inculcate field skills related to the domain subject in students. The syllabus of SEC will be partially skill oriented. Hence, teachers shall also impart practical training to students on the field skills embedded in the syllabus citing related real field situations.

Facilities requirements for teaching the Skill Enhancement Courses in Statistics:

All the 6 courses mentioned as Skill Enhancement Courses require a computer lab with at least 20 desk top systems since each student has to practice the methodology and acquire skills in producing output. This requires various software packages both standard (like Microsoft Excel, TORA/ LINDO/LINGO for Operations Research,) as well as open-source packages (like R and Python) to handle statistical analysis. Once these facilities are provided and training given on these lines, the outgoing students will fit into skill oriented jobs like business data analysts and Data Scientists.

Semester – wise Revised Syllabus under CBCS, 2020-21

Course Code:

Four-year B.A./B.Sc.

Domain Subject: **Statistics**

Course 6A: **OPERATIONS RESEARCH - I**

(Skill Enhancement Course(Elective), 05 Credits

Max. Marks: Theory :100 + Practicals: 50

Objective: The Objective of the paper is to introduce the basic concepts of Operational Research and linear programming to the students.

Learning Outcomes:

After learning this course, the student will be able

1. To know the scope of Operations Research
2. To link the OR techniques with business environment and life sciences
3. To convert real life problems into mathematical models
4. To find a solution to the problem in different cases
5. To inculcate logical thinking to find a solution to the problem

UNIT-I

Introduction of OR – Origin and development of OR – Nature and features of OR –Scientific Method in OR – Modeling in OR – Advantages and limitations of Models-General Solution methods of OR models – Applications of Operation Research. Linear programming problem (LPP) -Mathematical formulation of the problem - illustrations on Mathematical formulation of Linear programming of problem. Graphical solution of linear programming problems. Some exceptional cases - Alternative solutions, Unbounded solutions, non-existing feasible solutions by Graphical method.

UNIT-II

General linear programming Problem(GLP) – Definition and Matrix form of GLP problem, Slack variable, Surplus variable, unrestricted Variable, Standard form of LPP and Canonical form of LPP. Definitions of Solution, Basic Solution, Degenerate Solution, Basic feasible Solution and Optimum Basic Feasible Solution. Introduction to Simplex method and Computational procedure of simplex algorithm. Solving LPP by Simplex method (Maximization case and Minimization case)

UNIT-III

Artificial variable technique - Big-M method and Two-phase simplex method, Degeneracy in LPP and method to resolve degeneracy. Alternative solution, Unbounded solution, Non existing feasible solution and Solution of simultaneous equations by Simplex method.

UNIT-IV

Duality in Linear Programming –Concept of duality -Definition of Primal and Dual Problems, General rules for converting any primal into its Dual, Economic interpretation of duality, Relation between the solution of Primal and Dual problem(statements only). Using duality to solve primal problem. Dual Simplex Method.

UNIT-V

Post Optimal Analysis- Changes in cost Vector **C**, Changes in the Requirement Vector **b** and changes in the Coefficient Matrix **A**. Structural Changes in a LPP.

Reference Books:

1. S.D. Sharma, Operations Research, Kedar Nath Ram Nath & Co, Meerut.
2. Kanti Swarup, P.K.Gupta, Manmohn, Operations Research, Sultan Chand and sons, New Delhi.
3. J.K. Sharma, Operations Research and Application, Mc.Millan and Company, New Delhi.
4. GassS.I: Linear Programming. Mc Graw Hill.
5. HadlyG :Linear programming. Addison-Wesley.
6. Taha H.M: Operations Research: An Introduction : Mac Millan.

Practical/Lab to be performed on a computer using OR/Statistical packages

1. To solve Linear Programming Problem using Graphical Method with
 - (i) Unbounded solution
 - (ii) Infeasible solution
 - (iii) Alternative or multiple solutions.
2. Solution of LPP with simplex method.
3. Problem solving using Charne's M - method.
4. Problem solving using Two Phase method.
5. Illustration of following special cases in LPP using Simplex method
 - (i) Unrestricted variables
 - (ii) Unbounded solution
 - (iii) Infeasible solution
 - (iv) Alternative or multiple solutions.
6. Problems based on Principle of Duality.
7. Problems based on Dual simplex method.
8. Problems based on Post Optimal Analysis.

Four-year B.A./B.Sc.
Domain Subject: **Statistics**
Course 7A: **OPERATIONS RESEARCH - II**
(Skill Enhancement Course(Elective), 05 Credits
Max. Marks: Theory :100 + Practicals: 50

Objective: To enrich the knowledge of students with advanced techniques of linear programming problem along with real life applications.

Learning Outcomes:

After learning this course, the student will be able

1. To solve the problems in logistics
2. To find a solution for the problems having space constraints
3. To minimize the total elapsed time in an industry by efficient allocation of jobs to the suitable persons.
4. To find a solution for an adequate usage of human resources
5. To find the most plausible solutions in industries and agriculture when a random environment exists.

UNIT -I

Transportation Problem- Introduction, Mathematical formulation of Transportation problem. Definition of Initial Basic feasible solution of Transportation problem- North-West corner rule, Lowest cost entry method, Vogel's approximation method. Method of finding optimal solution- MODI method(U-V method). Degeneracy in transportation problem, Resolution of degeneracy, Unbalanced transportation problem. Maximization TP. Transshipment Problem.

UNIT-II

Assignment Problem -Introduction, Mathematical formulation of Assignment problem, Reduction theorem (statement only), Hungarian Method for solving Assignment problem, Unbalanced Assignment problem. The Traveling salesman problem, Formulation of Traveling salesman problem as an Assignment problem and Solution procedure.

UNIT-III

Sequencing problem: Introduction and assumptions of sequencing problem, Sequencing of n jobs and one machine problem. Johnson's algorithm for n jobs and two machines problem- problems with n-jobs on two machines, Gantt chart, algorithm for n jobs on three machines problem- problems with n- jobs on three machines, algorithm for n jobs on m machines problem, problems with n-jobs on m-machines. Graphical method for two jobs on m- machines.

UNIT-IV

Network Scheduling: Basic Components of a network, nodes and arcs, events and activities— Rules of Network construction – Time calculations in networks - Critical Path method (CPM) and PERT.

UNIT –V

Game Theory: Two-person zero-sum games. Pure and Mixed strategies. Maxmin and Minimax Principles - Saddle point and its existence. Games without Saddle point- Mixed strategies. Solution of 2×2 rectangular games. Graphical method of solving $2 \times n$ and $m \times 2$ games. Dominance Property. Matrix oddment method for $n \times n$ games. Only formulation of Linear Programming Problem for $m \times n$ games.

Reference Books:

1. S.D. Sharma, Operations Research, Kedar Nath Ram Nath & Co, Meerut.
2. Kanti Swarup, P.K.Gupta, Manmohn, Operations Research, Sultan Chand and sons, New Delhi.
3. J.K. Sharma, Operations Research and Application, Mc. Millan and Company, New Delhi.
4. Gass: Linear Programming. Mc Graw Hill.
5. Hadly : Linear programming. Addison-Wesley.
6. Taha : Operations Research: An Introduction : Mac Millan.
7. Dr.NVS Raju; Operations Research, SMS education,

Practical/Lab to be performed on a computer using OR/Statistical packages

1. IBFS of transportation problem by using North- West corner rule, Matrix minimum method and VAM
2. Optimum solution to balanced and unbalanced transportation problems by MODI method (both maximization and minimization cases)
3. Solution of Assignment problem using Hungarian method (both maximization and minimization cases),
4. Solution of sequencing problem—processing of n jobs through two machines
5. Solution of sequencing problem- processing of n jobs through three machines
6. To perform Project scheduling of a given project (Deterministic case-CPM).
7. To perform Project scheduling of a given project (Probabilistic case-PERT).
8. Graphical method of solving for $m \times 2$ and $2 \times n$ games.
9. Solution of $m \times n$ games by dominance rule.
10. Solution of $n \times n$ games by using matrix oddment method.
11. Linear programming method for solving $m \times n$ games.

Four-year B.A./B.Sc.
Domain Subject: **Statistics**
Course 6B: **Statistical Process and Quality Control**
(Skill Enhancement Course(Elective), 05 Credits
Max. Marks: Theory :100 + Practicals: 50

Course Objectives: To understand the concept of quality, process control and product control using control chart techniques and sampling inspection plan. To have an idea about quality management, quality circles, quality movement and standardizations for quality.

Learning Outcomes:

After learning this course, the student will be able

1. To define 'quality' in a scientific way
2. To differentiate between process control and product control
3. To speak about quality awareness in industry
4. To pave a path to an industry to meet the standards
5. To effectively implement various plans to control the quality standards at various stages of an industry.

Unit I

Meaning of quality, concept of total quality management (TQM) and six-sigma, ISO, comparison between TQM and Six Sigma, Meaning and purpose of Statistical Quality Control (SQC), Seven Process Control Tools of Statistical Quality Control (SQC) (i) Histogram (ii) Check Sheet, (iii) Pareto Diagram (iv) Cause and effect diagram (CED), (v) Defect concentration diagram (vi) Scatter Diagram (vii) Control chart. (Only introduction of 7 tools is expected).

Unit II

Statistical basis of Shewhart control charts, use of control charts. Interpretation of control charts, Control limits, Natural tolerance limits and specification limits. Chance causes and assignable causes of variation, justification for the use of 3-sigma limits for normal distribution, Criteria for detecting lack of control situations:

- (i) At least one point outside the control limits
- (ii) A run of seven or more points above or below central line.

Unit III

Control charts for Variables: Introduction and Construction of \bar{X} and R chart and Standard Deviation Chart when standards are specified and unspecified, corrective action if the process is out of statistical control.

Control charts for Attributes: Introduction and Construction of p chart, np chart, C Chart and U charts when standards are specified and unspecified, corrective action if the process is out of statistical control.

Unit IV

Acceptance Sampling for Attributes: Introduction, Concept of sampling inspection plan, Comparison between 100% inspection and sampling inspection. Procedures of acceptance sampling with rectification, Single sampling plan and double sampling plan.

Producer's risk and Consumer's risk, Operating characteristic (OC) curve, Acceptable Quality Level (AQL), Lot Tolerance Fraction Defective (LTFD) and Lot Tolerance Percent Defective (LTPD), Average Outgoing Quality (AOQ) and Average Outgoing Quality Limit (AOQL), AOQ curve, Average Sample Number (ASN), Average Total Inspection (ATI).

Unit V

Single Sampling Plan: Computation of probability of acceptance using Binomial and Poisson approximation, of AOQ and ATI. Graphical determination of AOQL, Determination of a single sampling plan by: a) lot quality approach b) average quality approach.

Double Sampling Plan: Evaluation of probability of acceptance using Poisson distribution, Structure of OC Curve, Derivation of AOQ, ASN and ATI (with complete inspection of second sample), Graphical determination of AOQL, Comparison of single sampling plan and double sample plan.

Text Books:

1. **Montgomery, D. C. (2008)**: Statistical Quality Control, 6thEdn., John Wiley, New York.
2. **Parimal Mukhopadhyay**: Applied Statistics, New Central Book Agency.
3. **Goon A.M., Gupta M.K. and Das Gupta B. (1986)**: Fundamentals of Statistics, Vol. II, World Press, Calcutta.
4. **S.C. Gupta and V.K. Kapoor**: Fundamentals of Applied Statistics – Chand publications.

References:

1. **R.C. Gupta**: Statistical Quality Control.
2. **Duncan A.J. (1974)**: Quality Control and Industrial Statistics, fourth edition D.B. Taraporewala Sons and Co. Pvt. Ltd., Mumbai.
3. **Grant, E. L. and Leavenworth (1980)**: Statistical Quality Control, fifth edition, McGraw Hill, New Delhi.

Practical/Lab to be performed on a computer using Statistical packages

1. Construction of \bar{X} and R Charts.
2. Construction of \bar{X} and σ Charts.
3. Construction of p Charts for fixed sample size.
4. Construction of p Charts for variable sample size.
5. Construction of np Charts.
6. Construction of C charts.
7. Construction of U charts.
8. Single sampling plan for attributes (OC Curve, Producer's and Consumer's risks, AOQ, AOQL, ATI).
9. Determination of single sampling plan by: a) lot quality approach b) average quality approach.
10. Double sampling plan for attributes (OC curve, AOQ, AOQL, ATI, ASN using Poisson distribution).

Four-year B.A./B.Sc.
Domain Subject: **Statistics**
Course 7B: **Computational Techniques and R Programming**
(Skill Enhancement Course (Elective), 05 Credits
Max. Marks: Theory :100 + Practicals: 50

Course Objectives: To learn the statistical analysis with the help of the statistical software R

Learning Outcomes:

After learning this course the student will be able

1. Understand the basic functioning of a computer
2. Acquire skills in handling business and organizational data using Excel
3. Perform simple analytics using Excel
4. Understand the power of R programming language
5. Handle various statistical issues using R language

Unit I

Computer basics: Basic applications of computer, components of computer system, Central Processing Unit (CPU), input and output units, computer memory and mass storage devices. Programming languages and their applications. Concept of files and folders. Software and types of software. Operating System like Windows and Linux.

Unit II

Data processing using spreadsheets – Data entry and editing features in Excel, copy, paste, paste special options, sort and filter options, auto sum, steps of finding average and standard deviation of data using statistical functions. Matrix operations like transpose, multiply and inverse using Excel functions. Simple graphs like bar chart, line chart and pie chart in Excel. Exporting Excel output to word processors like MS-Word and slide presentations like Power Point.

Unit III

Scatter diagram, fitting of straight line, polynomial and power curves using Excel – Reading R-square value and equation from the graph. Predicting future values using ‘forecast’ and ‘trend’ functions. Data Analysis Pak and its features. Performing Student’s t-test and one-way Analysis of Variance using Data Analysis Pak. P-value and its interpretation.

Unit IV

Programming with R: Introduction to R, Data types in R (numeric, logical, character, complex etc.), R objects: vector, matrix, array, list, data frame, factor, and time series. Arithmetic, logical and relational operators, explicit and implicit looping, functions and functional programming in R, Lexical scoping rules in R, benefits of Lexical scoping, other scoping rules, debugging facility in R. Few important mathematical, statistical and graphical functions in R.

Unit V

Descriptive Statistics with R software: Calculations with R software such as descriptive statistics, frequency distribution, Graphics and plots, statistical functions of central tendency, variation, skewness and kurtosis and illustration with examples.

Suggested Books

1. Chambers, J. (2008). Software for Data Analysis: Programming with R, Springer.
2. Crawley, M.J. (2017). The R Book, John Wiley & Sons.
3. Eckhouse, R.H. and Morris, L.R. (1975). Minicomputer Systems Organization, Programming and Applications, Prentice-Hall.
4. Matloff, N. (2011). The Art of R Programming, No Starch Press, Inc.
5. Peter N. (1986). Inside the IBM PC, Prentice-Hall Press.
6. Dr. Mark Gardener(2012): Beginning R The statistical Programming Languages, John Wiley & Sons.
7. K.V.S. Sarma (2010), Statistics Made Simple – Do it yourself on PC, 2nd Edition, Prentice Hall India
8. Sudha G. Purohit, SharadD.Gore, and ShailajaR.Deshmukh (2008), Statistics Using R, Narosa Publishing House, India.
9. Crawley, M.J. (2006). Statistics – An introduction using R. John Wiley London.
10. Purohit, S.G., Deshmukh, S.R. and Gore, S.D., (2015): Statistics using R, Alpha Science International.
11. Verzani, J., (2018): Using R for introductory statistics. CRC press.
12. Schumacker, R.E., (2014): Learning statistics using R. Sage Publications.
13. Michale J. Crawley (2009), THE R BOOK, John Wiley & Sons.

Practical/Lab to be performed on a computer using R Software

1. Construction of Bar Chart and Pie Chart using Excel
2. Fitting of straight-line using Excel
3. Calculating Matrix Inverse using Excel
4. One way ANOVA using Excel
5. Data visualization using R - frequency polygon, Ogives, Histogram.
6. Data visualization using R - simple and multiple bar diagram, pie chart.
7. Computation of Descriptive Statistics using R - Central Tendencies, Dispersions, Moments, Skewness and Kurtosis.
8. Computation of Karl Pearson's Coefficient of Correlation and Rank Correlation using R.
9. Construction of Control Charts for variables (\bar{X} , R and σ) charts using R.
10. Construction of Control Charts for attributes (p, np charts with fixed and varying sample sizes) using R.
11. Construction of Control Charts using R - C and U charts.

Four-year B.A./B.Sc.
Domain Subject: **Statistics**
Course 6C: **ECONOMETRICS**
(Skill Enhancement Course(Elective), 05 Credits
Max. Marks: Theory :100 + Practicals: 50

Learning Outcomes:

The course on econometrics will primarily focus on the use of statistical modeling and the relevant analyses to economic data problems. After learning this course the student will be able understand

1. various important econometric models and relevant model building concepts in econometrics
2. general linear models and estimation of inherent model parameters
3. multicollinearity, its detection and consequences and related inferential aspects
4. some advanced concepts of generalised least squares estimation, autocorrelation, its consequences, detection and strategy for reducing autocorrelation,
5. heteroscedasticity and its inherent concepts including its consequences,
6. some inferential aspects on heteroscedasticity,
7. practical aspects and real data illustration of the related problems.

UNIT-I

Basic Econometrics: Nature of econometrics and economic data, concept of econometrics, steps in empirical economic analysis, econometric model, importance of measurement in economics, the structure of econometric data, cross section, pooled cross section, time series and paired data.

UNIT-II

Models and Estimations: Simple regression models- two variable linear regression model, assumptions and estimation of parameters. Gauss Markoff theorem, OLS estimations, partial and multiple correlations coefficients. The general linear model assumptions, estimation and properties of estimators, BLUEs.

UNIT- III

Heteroscedastic disturbances : Tests of significance of estimators, R square and ANOVA. Concepts and consequences of heteroscedasticity. Tests and solutions of heteroscedasticity. Specification error, Errors of measurement.

UNIT-IV

Multicollinearity: The concept of multicollinearity and its consequences on econometric models, detection of multicollinearity. Measure of Multicollinearity – Variance Inflation Factor (VIF) and tolerance, formula and interpretation. Methods of reducing the influence of multicollinearity.

UNIT - V

Autocorrelation: Disturbance term (u) in econometric models and its assumptions, autocorrelated disturbances and their consequences on the model parameters, Detecting the presence of autocorrelation – hypothesis tests for autocorrelation - Durbin Watson test and its interpretation.

References:

1. Gujarati, D. and Sangeetha, S. (2007). Basic Econometrics, 4th Edition, McGraw Hill Companies.
2. Johnston, J. (1972). Econometric Methods, 2nd Edition, McGraw Hill International.
3. Koutsoyiannis, A. (2004). Theory of Econometrics, 2nd Edition, Palgrave Macmillan Limited.
4. Maddala, G.S. and Lahiri, K. (2009): Introduction to Econometrics, 4th Edition, John Wiley & Sons.

Practical/Lab to be performed on a computer using Statistical packages

1. Problems based on estimation of General linear model.
2. Testing of parameters of General linear model.
3. Forecasting of General linear model.
4. Problems concerning specification errors.
5. Problems related to consequences of Multicollinearity.
6. Diagnostics of Multicollinearity.
7. Problems related to consequences of Autocorrelation (AR(I)).
8. Diagnostics of Autocorrelation.
9. Estimation of problems of General linear model under Autocorrelation.
10. Problems related to consequences Heteroscedasticity.
11. Diagnostics of Heteroscedasticity.
12. Estimation of problems of General linear model under Heteroscedastic distance terms.

Four-year B.A./B.Sc.
Domain Subject: **Statistics**
Course 7C: **REGRESSION ANALYSIS**
(Skill Enhancement Course(Elective), 05 Credits
Max. Marks: Theory :100 + Practicals: 50

Learning Outcomes:

After learning this course the student will be able

1. To know about regression techniques, which are powerful tools in statistics,
2. To get an idea of Linear and Multiple Linear regression,
3. To learn about regression diagnostics, residual plots for visualization
4. To perform statistical tests of hypotheses on regression coefficients.
5. To study the structural stability of a regression model.
6. To learn the regression with qualitative independent and dependent variables by dummy variable technique.
7. To learn the selection of the best regression model.

Unit I

Simple Linear Regression: Simple Linear Regression Model. Least-Squares Estimation of the Parameters - Estimation of β_0 and β_1 , Properties of the Least-Squares Estimators and the Fitted Regression Model. Hypothesis Testing on the Slope and Intercept -Use of t Tests, Testing Significance of Regression and Analysis of Variance

Unit II

Multiple Linear Regression: Multiple linear regression: Multiple Linear Regression Model. Estimation of model parameters: Least-Squares Estimation of the Regression Coefficients, Properties of the Least-Squares Estimators. Concept of residual, Residual plots. Test for Significance of Individual Regression Coefficients, and subsets of coefficients. Concept of coefficient of determination.

Unit III

Regressions with Qualitative Independent Variables: Use of dummy variables to handle categorical independent variables in regression. Estimation of model parameters with dummy variables - Testing the structural stability of regression models, comparing the slopes of two regression models. Multiple linear regression with interaction effects.

Unit-IV

Regressions with Qualitative Dependent Variables: Binary logistic regression with several independent variables, estimation of coefficients, evaluating the Odds Ratio (OR) and its interpretation. The concept of Piecewise linear regression, The Logit, Probit and Tobit models and their applications.

Unit – V

Best Model: Selecting ‘Best’ regression model. All possible regressions – R^2 , Adjusted R^2 , MS_{Res} , Mallows’s statistic. Sequential selection of variables – criteria for including and eliminating a variable – forward selection, backward elimination and stepwise regression.

References:

1. Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining (2012), Introduction To Linear Regression Analysis, Fifth Edition, John Wiley & Sons
2. Draper, N. R. and Smith, H. (1998). Applied Regression Analysis. 3rd Edition. John Wiley.
3. Hosmer, D. W., Lemeshow, S. and Sturdivant R.X. (2013). Applied Logistic Regression, Wiley Blackwell.
4. Montgomery, D. C., Peck, E. A. and Vining, G. G. (2013). Introduction to Linear Regression Analysis. 5th Edition. Wiley.
5. Neter, J., Kutner, M. H., Nachtsheim, C.J. and Wasserman, W. (1996). Applied Linear Statistical Models, 4th Edition, Irwin USA.
6. Gujarati, D. and Sangeetha, S. (2007). Basic Econometrics, 4th Edition

Practical/Lab to be performed on a computer using Statistical packages

1. Least Squares estimates of slope and intercept
2. Plotting of two Regression Lines
3. Finding R-square value of Linear Models
4. Student’s t-test for regression coefficient
5. ANOVA for Multiple Linear Regression model
6. Selecting best regression model by R^2
7. Selecting best regression model by Adjusted R^2
8. Selecting best regression model by MS_{Res}
9. Selecting best regression model by Mallows’s statistic
10. Selecting best regression model by forward selection
11. Selecting best regression model by backward elimination.